

TD3 : Séries chronologiques

On considère une **série chronologique** Y_t pour $t = 1, \dots, T$:

- Y_t est la quantité observée
- T est le nombre total d'observations
- t le numéro de l'observation

On notera également :

- n le nombre de **périodes**
- p est le nombre de **saisons** par période

À partir des données des dernière années, on cherche à prédire une évolution future.

1 Vocabulaire

On considère le chiffre d'affaire d'une entreprise en k€ en fonction du mois et de l'année :

Chiffre d'affaire (en k€)	Année 1	Année 2	Année 3	Année 4
Janvier	22	22	23	23
Février	10	10	11	10
Mars	11	11	12	12
Avril	12	12	13	13
Mai	17	19	18	21
Juin	20	20	21	21
Juillet	17	17	18	20
Août	10	10	11	12
Septembre	21	20	21	22
Octobre	17	17	19	19
Novembre	15	15	16	16
Décembre	28	30	30	33

1. Quelle est la quantité observée ?

Y_t est

2. Combien y a-t-il d'observations au total ?

$T =$

3. Quel est le nombre de périodes ?

$n =$

4. Quel est le nombre de saisons par période ?

$p =$

2 Modèle additif ou modèle multiplicatif ?

5. Sur le fichier Excel, faire un nouveau tableau organisé comme indiqué ci-dessous :

Année	Mois	t	Y_t
01-année	01-janvier	1	22
01-année	02-février	2	10
...
01-année	12-décembre	12	28
02-année	01-janvier	13	22
02-année	02-février	14	10
...
04-année	12-décembre	48	33

6. Tracer sur un premier graphique Y_t en fonction de t . Penser à ajuster les axes pour que le graphique soit lisible. Que remarquez-vous visuellement ?

7. À l'aide d'**Insertion/Illustrations/Figures** tracer la droite qui passe par les maxima et la droite qui passe par les minima.

- **Modèle additif** : $Y_t = T_t + S_t + \epsilon_t$, les droites tracées à la question 7. sont **parallèles**
- **Modèle multiplicatif** : $Y_t = T_t \times S_t + \epsilon_t$, les droites tracées à la question 7. ne sont **pas parallèles**
où
 - T_t est la **tendance** (temps long)
 - S_t est la **saisonnalité** (temps court)
 - ϵ_t est le **bruit**

8. Sommes-nous en présence d'un modèle additif ou d'un modèle multiplicatif ? Justifier.

9. Tracer sur un second graphique Y_t en fonction du mois pour l'année 1 dans une couleur, puis sur le **même graphique**, superposer les données pour les années 2, 3 et 4. Penser à ajuster les axes pour que le graphique soit lisible.

Que remarquez-vous visuellement ?

10. On souhaite **prédire par le calcul** ce qu'il va se passer au cours de l'année 5 pour l'entreprise en terme de chiffre d'affaire. Revenez sur le tableau précédent et ajouter les lignes correspondant à une cinquième année. On laissera bien évidemment les cases Y_t pour $t \geq 49$ vides.

3 Saisonnalité

11. On définit la **moyenne mobile** MM_t sur l'observation des t :

$$MM_t = (0.5 \times Y_{t-6} + Y_{t-5} + \dots + Y_{t+5} + 0.5 \times Y_{t+6}) / 12$$

Ajouter et compléter une colonne MM_t à votre tableau.

Attention : les six premières cases et les six dernières cases de cette colonne ne peuvent pas être remplies !

12. Sur le premier graphique, tracer dans une autre couleur MM_t en fonction de t .

Que remarquez-vous ?

13. On pose $S_t = \frac{Y_t}{MM_t}$. Ajouter et remplir une colonne S_t .

Attention : les six premières cases et les six dernières cases de cette colonne ne peuvent pas être remplies !

14. Compléter le tableau suivant avec les valeurs de S_t correspondantes, puis donner la médiane pour chaque mois. Cette médiane sera noté S_t^* et sera dépendante du mois uniquement.

M \ A	Jan	Fév	Mar	Av	Mai	Juin	Juil	Août	Sep	Oct	Nov	Déc
Année 1												
Année 2												
Année 3												
Année 4												
médiane S_t^*												

15. Ajouter une colonne S_t^* au grand tableau et compléter la colonne avec les valeurs correspondantes.

4 Tendances

16. On appelle **corrigée des valeurs saisonnières** la quantité $CVS_t = \frac{Y_t}{S_t^*}$. Ajoutez une colonne CVS_t au grand tableau et compléter la colonne avec les valeurs correspondantes.

17. Sur un nouveau graphique, tracer CVS_t et fonction de t .

18. Effectuer une régression linéaire et donner la pente a et l'ordonnée à l'origine b

$$a =$$

$$b =$$

19. On pose $T_t = a \times t + b$. Ajouter et compléter une colonne T_t dans le grand tableau.

5 Prédictions

20. On pose $\hat{Y}_t = T_t \times S_t^*$. Ajouter et compléter une colonne \hat{Y}_t dans le grand tableau.
On obtient ainsi les prédictions pour l'année 5.
21. Sur un nouveau graphique, tracer Y_t et \hat{Y}_t en fonction de t .
On peut ainsi visualiser les prédictions pour l'année 5

6 Résidus

22. Les **résidus** $R_t = Y_t - Y_t^*$. Ajouter et compléter une colonne R_t dans le grand tableau.
23. Calculer l'erreur-type s
- $$s = \frac{1}{\sqrt{T-2}} \times \sqrt{(Y_1 - \hat{Y}_1)^2 + \dots + (Y_T - \hat{Y}_T)^2}$$
24. Les **résidus normalisés** sont définis par $Z_t = \frac{R_t}{s}$.
Ajouter et compléter une colonne Z_t dans le grand tableau.
25. Si X est une variable aléatoire suivant une loi normale centrée réduite, on définit Z_t^* comme le quartile tel que :
- $$P(X < Z_t^*) = \frac{t}{T}$$
- En utilisant **loi.normale.standard.inverse()**, ajouter et compléter une colonne Z_t^* dans le grand tableau.
Attention : la dernière case de cette colonne ne peut pas être remplie !
26. Sélectionner les colonnes Z_t et Z_t^* et trier les colonnes. Tracer Z_t^* en fonction de Z_t , puis effectuer une régression linéaire. Que peut-on déduire de la valeur de R^2 ?

7 Intervalle de confiance

27. Calculer \bar{t} la moyenne des t .
28. À l'aide de **somme.carres.ecarts()**, calculer C la somme des carrés des écarts à la moyenne \bar{t} .
29. W est une variable aléatoire suivant une loi de Student à $T - 2$ degrés de liberté.
En utilisant **loi.student.inverse.n()**, déterminer le quartile q tel que :
- $$P(W < q) = 0.975$$
30. Calculer $b_t = s \times q \times \sqrt{1 + \frac{1}{T} + \frac{(t-\bar{t})^2}{C}}$.
En déduire les intervalles de confiance I_t à 95% des \hat{Y}_t .